Published in final edited form as:
JAMA. 2005 February 16; 293(7): 799–809. doi:10.1001/jama.293.7.799.

Conversion Analysis for Mutation Detection in *MLH1* and *MSH2* in Patients With Colorectal Cancer

Graham Casey, PhD, **Noralane M. Lindor, MD**, **Nickolas Papadopoulos, PhD**, **Stephen N. Thibodeau, PhD**, **John Moskow, PhD**, **Scott Steelman, PhD**, **Carolyn H. Buzin, PhD**, **Steve S. Sommer, MD, PhD**, **Christine E. Collins, BS**, **Malinda Butz, BS**, **Melyssa Aronson, MS**, **Steven Gallinger, MD**, **Melissa A. Barker, MSc**, **Joanne P. Young, PhD**, **Jeremy R. Jass, MD**, **John L. Hopper, PhD**, **Anh Diep, BS**, **Bharati Bapat, PhD**, **Michael Salem, MD**, **Daniela Seminara, PhD, MPH**, **Robert Haile, PhD**, and **the Colon Cancer Family Registry**
Department of Cancer Biology, Cleveland Clinic Lerner College of Medicine, Cleveland, Ohio (Dr Casey); Departments of Medical Genetics (Dr Lindor) and Laboratory Medicine and Pathology (Dr Thibodeau and Ms Butz), Mayo Foundation, Rochester, Minn; GMP Genetics, Waltham, Mass (Drs Papadopoulos, Moskow, Steelman, and Salem and Ms Collins); Clinical Molecular Diagnostic Laboratory, City of Hope National Medical Center, Duarte, Calif (Drs Buzin and Sommer); Departments of Surgery and Pathology, Mt Sinai Hospital, Toronto, Ontario (Ms Aronson and Drs Gallinger and Bapat); Cojoint Gastroenterology Laboratory, Bancroft Centre, Herston, Australia (Ms Barker and Dr Young); Department of Pathology, McGill University, Montreal, Quebec (Dr Jass); Centre for Genetic Epidemiology, University of Melbourne, Melbourne, Australia (Dr Hopper); Department of Preventive Medicine, University of Southern California–Norris Cancer Center, Los Angeles (Ms Diep and Dr Haile); National Cancer Institute, National Institutes of Health, Bethesda, Md (Dr Seminara).

Abstract

Corresponding Author: Graham Casey, PhD, Department of Cancer Biology, ND50, Cleveland Clinic Lerner College of Medicine, 9500 Euclid Ave, Cleveland, OH 44195 (caseyg@ccf.org)..

Author Contributions: Dr Casey had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Casey, Lindor, Papadopoulos, Thibodeau, Moskow, Steelman, Sommer, Jass, Hopper, Diep, Bapat, Salem, Seminara, Haile.
Acquisition of data: Casey, Lindor, Papadopoulos, Thibodeau, Moskow, Steelman, Buzin, Sommer, Collins, Butz, Aronson, Gallinger, Hopper, Diep, Bapat.
Analysis and interpretation of data: Casey, Lindor, Papadopoulos, Thibodeau, Moskow, Steelman, Buzin, Sommer, Butz, Barker, Young, Hopper, Bapat, Haile.
Drafting of the manuscript: Casey, Lindor, Papadopoulos, Moskow, Steelman, Hopper, Haile.
Critical revision of the manuscript for important intellectual content: Casey, Lindor, Papadopoulos, Thibodeau, Steelman, Buzin, Sommer, Collins, Butz, Aronson, Gallinger, Barker, Young, Jass, Hopper, Diep, Bapat, Salem, Seminara.
Statistical expertise: Hopper, Haile.
Obtained funding: Papadopoulos, Gallinger, Jass, Hopper, Salem, Seminara.
Administrative, technical, or material support: Casey, Lindor, Papadopoulos, Thibodeau, Moskow, Steelman, Collins, Butz, Gallinger, Barker, Young, Diep, Salem, Seminara.
Study supervision: Casey, Papadopoulos, Steelman, Buzin, Sommer, Aronson, Diep, Salem.

Financial Disclosures: None reported.

Role of the Sponsor: GMP Genetics Inc had input in the design of the study. GMP Genetics generated monoallelic hybrids for chromosome 2 and 3 from all live cell lines included in this study and performed the mutation analysis on nucleic acids isolated from these cells.

Disclaimer: This article does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Cancer Family Registries, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government or the Cancer Family Registry.

Context—The accurate identification and interpretation of germline mutations in mismatch repair genes in colorectal cancer cases is critical for clinical management. Current data suggest that mismatch repair mutations are highly heterogeneous and that many mutations are not detected when conventional DNA sequencing alone is used.

Objective—To evaluate the potential of conversion analysis compared with DNA sequencing alone to detect heterogeneous germline mutations in *MLH1*, *MSH2*, and *MSH6* in colorectal cancer patients.

Design, Setting, and Participants—Multicenter study with patients who participate in the Colon Cancer Family Registry. Mutation analyses were performed in participant samples determined to have a high probability of carrying mismatch repair germline mutations. Samples from a total of 64 hereditary nonpolyposis colorectal cancer cases, 8 hereditary nonpolyposis colorectal cancer–like cases, and 17 cases diagnosed prior to age 50 years were analyzed from June 2002 to June 2003.

Main Outcome Measures—Classification of family members as carriers or noncarriers of germline mutations in *MLH1*, *MSH2*, or *MSH6*; mutation data from conversion analysis compared with genomic DNA sequencing.

Results—Genomic DNA sequencing identified 28 likely deleterious exon mutations, 4 in-frame deletion mutations, 16 missense changes, and 22 putative splice site mutations. Conversion analysis identified all mutations detected by genomic DNA sequencing—plus an additional exon mutation, 12 large genomic deletions, and 1 exon duplication mutation—yielding an increase of 33% (14/42) in diagnostic yield of deleterious mutations. Conversion analysis also showed that 4 of 16 missense changes resulted in exon skipping in transcripts and that 17 of 22 putative splice site mutations affected splicing or mRNA transcript stability. Conversion analysis provided an increase of 56% (35/63) in the diagnostic yield of genetic testing compared with genomic DNA sequencing alone.

Conclusions—The data confirm the heterogeneity of mismatch repair mutations and reveal that many mutations in colorectal cancer cases would be missed using conventional genomic DNA sequencing alone. Conversion analysis substantially increases the diagnostic yield of genetic testing for mismatch repair mutations in patients diagnosed as having colorectal cancer.

HEREDITARY NONPOLYPOSIS COlorectal cancer (HNPCC) is a clinically heterogeneous disease that has historically been diagnosed based on family history (Amsterdam and Bethesda criteria).[1-3] Based on family history, approximately 70% of HNPCC cases and a proportion of cases that do not fit these criteria can be accounted for by mutations in any one of several genes involved in DNA mismatch repair. In those cases with defective mismatch repair, approximately 95% have alterations in 1 of 3 of the mismatch repair genes *MLH1*, *MSH2*, and *MSH6*, with a smaller proportion attributable to mutations in other mismatch repair genes.[4-6] Identification of a mutation may prompt genetic counseling, screening, and surveillance of relatives to reduce morbidity and mortality. It has been proposed that screening of all patients with colorectal cancer for mismatch repair gene mutations is both feasible and desirable.[5]

The accurate identification and interpretation of mismatch repair mutation carriers is essential for clinical management of colorectal cancer patients and for scientific studies in which the mutation status of participants is an important variable. Currently, most genetic testing is performed by genomic DNA sequence analysis, but certain classes of gene mutations are not detected using this approach, particularly large genomic deletions, genomic rearrangements, and loss of expression mutations.[7-9] Studies suggest that large genomic deletions may account for a substantial proportion of HNPCC cases in the United States.[10] Furthermore, whereas conventional genomic DNA sequencing methods may identify putative mutations in splice-site regions, additional studies are required to establish their pathogenic effect.[8]

The heterogeneous nature of mismatch repair mutations in colorectal cancer has not been well characterized, but there is evidence that large genomic mutations may account for a substantial proportion of cases.[7,9] Recent studies have suggested that conversion analysis, in which alleles

are separated in hybrids prior to mutation screening, represents a more sensitive mutation detection method than conventional DNA sequencing for identifying such mutations.[7,9,11,12] To date, only a limited number of studies have used conversion analysis to identify mutations in *MLH1* and *MSH2* in individuals for whom DNA sequencing failed to detect any mutations.[7,9] There have been no studies comparing the relative accuracy and specificity of conversion analysis with other mutation testing methods in a rigorous way, which is essential if the method is to be widely used clinically and in research.

In this study, we performed a blinded comparison of conventional DNA sequencing and conversion analysis to identify mutations in *MLH1*, *MSH2*, and *MSH6* in 89 colorectal cancer cases. These cases had a high probability of carrying a mutation in mismatch repair genes. We sought to define the full complement of mismatch repair mutations and to provide a strategy for the development of a comprehensive test for the identification of germline mismatch repair mutation carriers.

METHODS

Patient Recruitment

Participants were recruited through the Colon Cancer Family Registry, which is supported by the National Cancer Institute. The Colon Cancer Family Registry Consortium was initiated in 1997 and is dedicated to the establishment of a comprehensive collaborative infrastructure for interdisciplinary studies in the genetic epidemiology of colorectal cancer. The participating centers of the Colon Cancer Family Registry are the University of Hawaii (Honolulu); Fred Hutchinson Cancer Research Center (Seattle, Wash); Mayo Clinic (Rochester, Minn); University of Southern California (Los Angeles); University of Queensland (Brisbane, Australia); University of Melbourne (Melbourne, Australia); Cancer Care Ontario (Toronto). The 6 registries (and collaborating institutions) use standardized instruments and protocols to collect family history information, epidemiological and clinical data, screening behavior, and related biological specimens. Quality-control measures were in place throughout the collection, processing, and storing of data and samples. All study participants provided written consent and institutional review board approval for this study was obtained from all participating centers. Additional information about the Colon Cancer Family Registry can be found at http://www.cfr.epi.uci.edu/.

Participants were entered into the study through the Mayo Clinic, the University of Southern California Consortium, Cancer Care Ontario, and the University of Queensland. Participants had a prior diagnosis of colorectal cancer and an available Epstein-Barr virus–transformed cell line. A *participant* was defined as an individual who belonged to a family that met the Amsterdam 1 criteria for HNPCC; had at least 2 first- or second-degree relatives with colorectal cancer or 1 relative with endometrial cancer and at least 1 other relative with colorectal cancer (who did not meet Amsterdam 1 criteria and were referred to as HNPCC-like); or was diagnosed with colorectal cancer prior to age 50 years but did not meet Amsterdam 1 criteria and was not HNPCC-like. Importantly, in addition to these criteria, the majority of cases (85 of 89) were selected because they also had prior evidence of a defect in mismatch repair due to having either a tumor with high microsatellite instability or loss of expression of a mismatch repair protein by immunohistochemistry. Sixty-four HNPCC cases, 8 HNPCC-like, and 17 colorectal cancer cases diagnosed prior to age 50 years were entered into the study. The identity of the participant samples, family history of cases, and data on microsatellite instability or immunohistochemistry were blinded until the end of the study.

Microsatellite Instability Testing

Each center conducted its own microsatellite instability testing according to a defined protocol involving testing tumors for instability at 10 specific loci (BAT25, BAT26, ACTC, D5S346, D17S250, MYCL, BAT40, BAT34C4, D18S55, D10S197).[13] This microsatellite instability marker set included the National Cancer Institute recommended markers.14 A tumor was defined as having high microsatellite instability if instability was seen in more than 30% of the markers tested.

Immunohistochemical Analysis

Detailed immunohistochemistry staining methods have been described elsewhere.[13] Briefly, 5-μm tissue sections from formalin-fixed, paraffin-embedded tissue were stained using the avidin-biotin complex method of Ventana Medical Systems (Tucson, Ariz) (buffer kit and DAB detection kit, BioTek Solutions, Carpenteria, Calif) and using the Tech Mate 500 (Ventana) automated immunohistochemical stain. Staining was performed using antibodies to MLH1 (clone G168-728, 1/250; Pharmingen, San Diego, Calif), MSH2 (clone FE11, 1/50; Oncogene Research Products, Cambridge, Mass), MSH6 (clone 44, 1/500; Transductions Laboratories, Lexington, Ky), and PMS2 (clone A16-4, 1/100; Pharmingen). When tumor cells showed an absence of nuclear staining in the presence of positive staining in surrounding cells, absence of protein expression was interpreted.[13] Not all cases were screened for MSH6 and PMS2 protein staining.

Mutation Analyses

DNA Sequencing—DNA sequencing was performed on all cases for mutations in *MLH1* and *MSH2*. *MSH6* sequencing was performed only on those cases (n=23) that were negative for deleterious mutations in *MLH1* or *MSH2*. Genomic DNA was isolated from lymphoblastoid cell lines from each participant using a blood kit (Qiagen Inc, Valencia, Calif). Samples were coded and submitted to a clinical molecular diagnostic laboratory (City of Hope National Medical Center, Duarte, Calif) for *MLH1* and *MSH2* sequencing and to the molecular genetics laboratory (Mayo Clinic) for *MSH6* sequencing.

Polymerase chain reaction (PCR) amplification was performed on genomic DNA for all exons and adjacent intronic splice regions (19 fragments for *MLH1*, 16 for *MSH2*, and 16 for *MSH6*). All gene segments were sequenced in both directions using an ABI 377 or ABI 3100 automated DNA sequencer (Applied Biosystems, Foster City, Calif). Sequence changes were detected using Sequencher software (Gene Codes, Ann Arbor, Mich); all chromatograms were analyzed by at least 2 individuals. Mutations were confirmed by repeating the PCR amplification reaction and sequencing. All sequencing primers are available on request.

Conversion Analysis—Lymphoblastoid cell lines from all participants were coded and submitted to GMP Genetics (Waltham, Mass) for conversion analysis.7,15 Hybrid cell lines were generated following conversion technology protocols by fusing lymphoblastoid cells from participants with E2 mouse cells. The E2 cell line is an immortal mouse embryonic cell line. This cell line and properties are described.7 E2 cells were mixed with lymphoblastoid cells, washed in Hanks balanced salt solution, resuspended in fusion media, and transferred to a cuvette. Following electrofusion at 280V, the cells were plated in plastic tissue culture multiwell plates and grown in a Dulbecco modified eagle medium with 10% fetal bovine serum. Fusions were performed on a BTX electro cell manipulator (Genetronic, San Diego, Calif). An electro cell manipulator is an instrument that generates an electric current that facilitates cell fusion.

During the next day, the cells were fed (using a Dulbecco modified eagle medium) 10% fetal bovine serum supplemented with 1× hypoxanthine-aminopterin-thymidine and 0.5 mg/mL of

G418 (selection has been described7). Two weeks later, approximately 42 hybrid clones were further expanded. Cells were trypsinized and an aliquot was used for the isolation of DNA with reagents and protocols from blood kits (Qiagen).

A minimum of 2 hybrids for each allele of chromosome 2 and chromosome 3 were selected for isolation of RNA and further analysis. GMP Genetics Inc generated at least 2 monoallelic hybrids per allele for chromosome 2 and 3 from all live cell lines included in this study. Hybrids that contained human chromosomes 2 (for *MSH2* and *MSH6*) or 3 (for *MLH1*) were identified by genotyping DNA prepared from the hybrids using 16 highly polymorphic microsatellite DNA markers—8 each on human chromosomes 2 or 3. Hybrids containing chromosome 2 were identified by genotyping with the markers D2S2211, D2S162, D2S367, D2S337, D2S117, D2S325, D2S2382, and D2S206. Hybrids containing chromosome 3 were identified by genotyping with the markers D3S1297, D3S1304, D3S2338, D3S1289, D3S1614, D3S1565, D3S1262, and D3S1580. The markers used for genotyping were derived from the linkage mapping set (version 2.5, MD10; Applied Biosystems). Polymerase chain reaction amplification of the DNA markers was performed with Taq and reaction buffer (Invitrogen, Carlsbad, Calif). The PCR products were fractionated by capillary electrophoresis using an ABI 3100 automated DNA sequencer. Total RNA was isolated using the RNeasy Kit (Qiagen) and cDNA generated using SuperScript II reverse transcriptase (Invitrogen). Both positive and negative reverse transcriptase cDNA reactions were performed from each RNA source.

Conversion analysis combines the separation of alleles into hybrids along with analysis of cDNA sequence changes and effects on mRNA expression. Changes in mRNA transcript size or levels of mRNA expression of *MSH2* and *MLH1* genes were determined by reverse transciptase PCR from hybrids containing chromosomes 2 or 3. The coding region of each gene was amplified by PCR with overlapping fragments. Each fragment was amplified by PCR twice using 2 independent PCR reactions. To ensure the integrity of cDNA and to verify the presence of human allele, PCR fragments of other human genes of equivalent length to *MSH2*, *MSH6*, or *MLH1* were amplified. The PCR fragments were amplified from the same cDNA source. Specifically, human T-cell leukemia virus–enhanced factor (HTLF) was used as a control on hybrids with chromosome 2 alleles and human transferrin receptor (TFRC) on hybrids with chromosome 3 alleles. Conversion analysis of *MSH6* was performed only on those cases negative for deleterious mutations in *MLH1* or *MSH2*.

The primer pairs used to amplify the 2 fragments of *MSH2* were MSH2-6 (5′-GCGCATTTTCTTCAACCAGG-3′) and MSH2-18 (5′-TAATCTGTTTGCCAGGGTCC -3′) at an annealing temperature of 65°C and MSH2-20 (5′-CTGACTTCTCCAAGTTTCAGG -3′) and MSH2-4 (5 ′ -TGGGCACTGACAGTTAACAC-3′) at an annealing temperature of 60°C.

The primer pairs used to amplify the 2 fragments of *MLH1* were MLH1-14 (5′-CACTTCCGTTGAGCATCTAG-3′) and MLH1-2 (5′-GCTGCAGAAATGCATCAAGC-3′) at an annealing temperature of 60°C and MLH1-17 (5′-CAGCACATCGAGAGCAAGC -3′) and MLH1-18 (5′-ATCACACTTTGATACAACACTTTG -3′) at an annealing temperature of 63°C.

The primer pairs used to amplify the 6 fragments of *MSH6* had either M13 forward (5′ primers) or M13 reverse (3′ primers) to facilitate sequencing after amplification. The 6 pairs were: MSH6-31(5′- GGGCCTTGCCGGCTGTC GGT-3′) and MSH6-4 (5′-CTAGATCCTTGTGTCTTAGGCTGTACTTCC-3′); MSH6-6 (5′-CTCAGAGCCAGAGAAGAGGAAGA AGAGATG-3′) and MSH6-8 (5′-CTGACTCCAATAAGAGCATCCATGTGTACAG-3”); MSH6-10 (5′-AGTCTCAGAACTTTGATCTTGTC ATCTGTTAC-3′) and MSH6-11 (5′-CAATAAGGCATTTTTTGAGGTAGAAGACAC-3′); MSH6-14 (5′-

GGGAAAAGCTAAGTGATGGCAT TGGGG-3′) and MSH6-15 (5′-TGGTCAAAGGCTGTATCCCATCGG-3′); MSH6-18 (5′-GGTTTTAAGTCTAAAATCCTTAAGCAGGTC-3′) and MSH6-19 (5′-CAGCTAATAAGCCAGCCTGTCTCATAAGC-3′); and MSH6-21 (5′-ATGACATTCTAATAGGCTGTGAGGAAGAG-3 ′) and MSH6-24 (5′-GTTGTCTGAATTTACCACCTTTGGTCAG-3′). The annealing temperature was 69°C for all 6 fragments.

The primer pairs used to amplify TFRC were TFRC-1 (5′-ATTCTGCTCGTGGAGACTTC -3′) and TFRC-3 (5′-CTTATCTGGTCAGTGCTCGC -3′) at an annealing temperature of 63° C. The primer pairs used to amplify *HTLF* were HTLF-3 (5 ′ -GACTCCAGATAAGAGAGCTG -3′) and HTLF-4 (5′-TTAGTATCCCTTCCCTACCC -3′) at an annealing temperature of 63°C.

The PCR reaction conditions were 10.0 μL of cDNA; 0.2 μL of 50 μM for primer 1; 0.2 μL of 50 μM for primer 2; 5.0 μL of 10 × PCR buffer (included with Taq polymerase); 0.25 μL of 100 μM of nucleotide mix (Invitrogen); 1.5 μL of 50 μM of magnesium chloride (included with Taq polymerase); 1.0 μL of platinum TAQ DNA polymerase (Invitrogen); and water to 50.0 μL. The PCR cycle conditions were 94°C (1 cycle) for 3 minutes, followed by 94°C for 30 seconds; primer-specific annealing temperature for 30 seconds; 72°C for 2 minutes (35 cycles); and 72°C for 5 minutes (1 cycle). The amplified fragments were resolved by running between 7 and 9 μL of the PCR reaction on a 1% agarose gel.

Sequencing of cDNA Products—The PCR fragments were purified using the AMPure purification system (Agen-court, Beverly, Mass). Sequencing reactions were performed using BigDye terminator cycle sequencing kit (version 3.1, Applied Biosystems). Sequencing products were resolved using an Aurora DNA sequencer (Spectrumedix Corp, State College, Pa). Sequencing data were analyzed using the Laser-Gene (DNASTAR Inc, Madison, Wis) and the Mutation Surveyor (Soft Genetics Inc, State College, Pa).

Southern Blot Analysis of Large Genomic Deletions—This method has been described in detail elsewhere.[16] Genomic DNA from each patient was digested separately with 3 restriction endonucleases, EcoR I, BglII, and Hind III. Each individual digested genomic DNA (2.5 mg) was then loaded onto a 0.8% agarose gel for overnight electrophoresis at 55 V. After standard capillary gel transfer to the hybridization membrane (approximately 10 ng for each of the purified probes; 3 probes for *hMSH2* capable of identify individual exons) was radioactively labeled with [a-32P]-dCTP using the high prime kit (Roche, Basel, Switzerland). The radioactive probes were added to 20 mL of hybridization solution at a concentration of approximately 1×10^6 counts per minute. Membranes were placed in the probe-hybridization solution and hybridization took place overnight at 45°C. Following hybridization, the membranes were washed 3 times in 2 × sodium chloride sodium citrate buffer and 0.1% sodium dodecyl sulfate at 60°C for 30 minutes; and then once in 0.2 × sodium chloride sodium citrate buffer and 0.1% sodium dodecyl sulfate at 60°C for 30 minutes. The radioactive membranes were then exposed to PhosphorImager screens (Amersham Biosciences, Piscataway, NJ). Following exposure, the PhosphorImager screens were scanned and the results were analyzed using ImageQuant software (version 5.0, Amersham Biosciences).

RESULTS

Cases were selected for study based on a number of clinical characteristics and the presence of defective mismatch repair either by microsatellite instability testing or immunohistochemistry, or both. Summaries of cancer family histories for each of the 3 groups are provided in Table 1 and mutation data are summarized in Table 2. Mutations were

considered pathogenic if the change met any of the following criteria: a frameshift mutation that would be predicted to result in a truncated protein; nonsense mutations; missense mutations if additional mRNA expression data revealed aberrant splicing or exon skipping; splice site mutations if additional data revealed aberrant splicing; large genomic deletions that removed at least 1 exon; or duplication of exons. Four in-frame deletions and missense mutations of unknown clinical significance were not classified as deleterious because additional data are needed.

Detailed mutation data for the 64 colorectal cancer cases from HNPCC Amsterdam 1 Criteria Families appear in Table 3 and Table 4. Table 5 presents mutation data for the 8 colorectal cases from HNPCC-like families. Table 6 presents mutation data for the 17 colorectal cancer cases who were diagnosed prior to age 50 years. Conventional genomic DNA sequence analyses identified 28 pathogenic coding domain mutations, 16 missense mutations of unknown clinical significance, 4 in-frame deletion mutations, and 22 mutations in splice-site regions within introns. Conversion analysis identified all 28 likely pathogenic coding domain mutations, plus 14 additional pathogenic mutations, including 1 exon mutation, 12 large genomic deletions, and 1 exon duplication. This represents an increase of 33% (14/42) in the number of likely pathogenic mutations detected by conversion analysis compared with those detected by conventional DNA sequencing. Further studies are needed to determine whether any of the in-frame deletion mutations are pathogenic.

To confirm that conversion analysis correctly identified large genomic deletions, we performed Southern blot analyses on a subset of 5 cases with this type of mutation (those cases are identification numbers 3, 16, 56, 62, and 94). Southern blot and conversion analysis data were consistent for all 5 cases.

Using reverse transcriptase PCR analysis of gene expression in hybrids, conversion analysis also provided evidence of a likely pathogenic role for 4 of the 14 missense mutations that could not be interpreted on the basis of the conventional genomic DNA sequencing alone. All 4 mutations are in coding regions adjacent to splice sites of the deleted exons. One mutation (*MLH1* 793C>T) identified in 2 different families (identification numbers 28 and 29) was located 3 base pairs from the 5′ end of exon 10 and was associated with skipping of exons 9 and 10 in cDNA. One missense mutation (*MSH2* 1660A>G in identification number 57) was located 2 base pairs from the 3′ end of exon 10 and was associated with skipping of exon 10 in cDNA. A fourth mutation (*MLH1* 883A>G) in identification number 50 was located 2 base pairs from the 3′ end of exon 10 and was associated with skipping of exons 9 and 10 in cDNA.

Conversion analysis studies also clarified the pathogenic effect of the 22 mutations within intron splice site regions (IVS+1 to IVS+5, IVS–2, and IVS–8). With the exception of the 5 *MLH1* IVS11-8 mutations and 2 *MLH1* IVS07-2 mutations, all the splice-site mutations affected splicing. The 2 cases with the *MLH1* IVS07-2 mutation showed loss of mRNA expression of the corresponding *MLH1* transcript. This change (*MLH1* IVS07-2) has been reported twice in the International Collaborative Group on Hereditary Non-Polypsis Colorectal Cancer (INSIGHT)[6] mutation database and would be predicted to result in a splicing defect. Both cases showed loss of mRNA transcript expression, suggesting that this mutation may result in unstable mRNA. Further studies are needed to confirm this finding. The overall diagnostic yield of detecting clinically relevant mutations using conversion analysis compared with conventional DNA sequencing was 56% (35/63).

COMMENT

In this study of 89 colorectal cancer patients, who were selected because of high probabilities of carrying a mutation in a mismatch repair gene due to their family history, age at diagnosis,

microsatellite instability, and/or loss of MLH1, MSH2, or MSH6 protein expression of their tumors, we identified likely pathogenic mutations in *MLH1, MSH2*, or *MSH6* in 63 (71%) of the 89 cases. Among the 3 groups with defective mismatch repair evaluated, likely deleterious mutations were identified in 77% (49/64) of the Amsterdam I criteria HNPCC cases, 88% (7/8) of HNPCC-like cases, and 41% (7/17) of colorectal cancer cases diagnosed prior to age 50 years.

Overall, we identified 29 likely pathogenic coding domain mutations, 17 mutations that affected splicing or mRNA transcript stability, 12 large genomic deletions, 1 exon duplication, 4 in-frame deletion mutations, and 12 missense mutations of unknown clinical significance. We do not report the 4 in-frame deletions as deleterious because additional studies are required to confirm their pathogenic effect.

Together, these data imply that the great majority of HNPCC and HNPCC-like colorectal cancer cases can be attributed to germline mutations in *MLH1* or *MSH2* when cases are preselected on the basis of tumor characteristics for harboring a likely mismatch repair defect. This frequency is higher than what would be anticipated from testing clinically selected cases based on family history alone. Our finding of 1 *MSH6* mutation carrier family in this population is consistent with *MSH6* mutations accounting for only a low percentage of colorectal cancer cases.[4-6] Note that colorectal cancer cases in this study were selected based on tumor microsatellite instability status and loss of *MLH1* or *MSH2* staining only.

Mutations in *MLH1* and *MSH2* (34% and 42%, respectively) accounted for a similar proportion of HNPCC Amsterdam 1 cases. This frequency is consistent with that reported by Wagner et al[9] (42% and 41% for MLH1 and MSH2 mutations, respectively) in their study of 49 Amsterdam 1 criteria HNPCC families and 10 HNPCC-like families. In contrast, we found that the majority of mutations identified in our HNPCC-like cases and colorectal cancer cases diagnosed prior to age 50 years were in the *MSH2* gene (11 in *MSH2*, 2 in *MLH1*, and 1 in *MSH6*), suggesting a greater variability in family history presentation in *MSH2* than *MLH1* mutation-related cases.

The high number of large genomic deletions seen in our study is consistent with the data from the study by Wagner et al,9 reporting a high frequency of these mutations in 49 Amsterdam 1 criteria HNPCC families and 10 HNPCC-like families. However, in the study by Wagner et al,[9] 50% (7/14) of the cases with large genomic alterations had the same founder mutation (deletion of exons 1-6 of *MSH2*), with the majority of cases belonging to a single-extended lineage arising from a common European ancestor. This mutation was found only once in our study population in a HNPCC-like family from North America and the relationship between this case and the extended family described by Wagner et al[9] and others[17] is not known. Our data imply that not only do large genomic deletions occur frequently in colorectal cancer cases, but also that this type of mutation is highly diverse.

The highly heterogeneous nature of germline mismatch repair mutations in HNPCC and other colorectal cancer patients presents serious challenges for accurate genetic testing. We have shown that mutation testing using genomic DNA sequencing alone would result in a high frequency of false-negatives within samples chosen because they were highly likely to carry a mutation. Additional analytical approaches are required to detect all of the mutations likely to occur in HNPCC cases with defective mismatch repair.

Conversion analysis in combination with cDNA sequencing and mRNA expression analysis offers a comprehensive approach for the detection of mismatch repair mutations that occur in HNPCC. Conversion analysis has been adapted to use blood samples (the main clinical material for a reference laboratory) rather than lymphoblastoid cell lines and has been shown to have high efficiency in generating hybrids, have a high capacity, and a turnaround time that is

acceptable to a reference laboratory. Furthermore, the results of this study and previously published work indicated that conversion analysis (conversion technology coupled with analysis of cDNA) can be used as a clinical platform for mutation screening of HNPCC and other diseases.

Mutations that were not detected by DNA sequencing were predominantly large genomic deletions that would be masked due to the presence of the remaining normal allele. The separation of alleles through conversion analysis allowed for the unmasking and detection of these mutations and also provided important information for interpreting the clinical significance (ie, the pathogenic nature) of both missense mutations and putative splice-site mutations. It should be noted that genomic DNA sequencing when used in combination with other analytical approaches, such as Southern blotting, multiplex ligation-dependent probe amplification,[18-21] or cDNA sequencing and reverse transcriptase PCR expression analysis, has the potential to provide the same information as conversion analysis. These other approaches, however, were not the subject of this analysis. Our study confirms the highly heterogeneous nature of mismatch repair mutations in colorectal cancer cases. Any mutation testing strategy that is adopted must take this heterogeneity into account. Our study warrants additional studies comparing conversion analysis with DNA sequencing used in combination with Southern blotting and multiplex ligation–dependent probe amplification.

We found the lowest frequency of mismatch repair mutations (41%) in the 17 colorectal cancer cases diagnosed as having tumors with high microsatellite instability prior to age 50 years without an HNPCC-like family history. Nevertheless, a mismatch repair gene mutation cannot be altogether excluded. Our findings are similar to those of Liu et al[22] who found pathogenic mutations in 5 (42%) of 12 participants with high microsatellite instability colorectal cancer who were diagnosed at age 35 years or younger. These data emphasize the importance of high microsatellite instability status as a marker for HNPCC in younger persons with no family history, especially if he/she shows a loss of protein expression. Such cases are sometimes erroneously dismissed as “sporadic” colorectal cancer, with the implication that they are not present in individuals with a predisposing germline mutation.23

We have previously reported on the correlation between microsatellite instability and immunohistochemistry results,[13] and the current study supports these findings. For those cases with likely deleterious mutations and available immunohistochemistry staining, data were consistent for 51 (96%) of 53 HNPCC and HNPCC-like cases. In this multicenter study, we were unable to reevaluate discordant immunohistochemistry and mutation data. Overall, this study supports the use of immunohistochemistry as a rapid, reasonably sensitive, and specific tool for triaging a specific mismatch repair gene for germline testing.

In conclusion, we have shown that DNA sequencing alone is not sufficiently sensitive to detect the types of mutations in *MLH1* and *MSH2* genes found in colorectal cancer cases. Conversion analysis provided a 33% improvement in the detection of mismatch repair mutations in 89 colorectal cancer cases selected as highly likely to have a mutation. The overall increase in clinically important information provided by conversion analysis was 56% (35/63). These results have important implications for genetic testing of individuals for both clinical and research purposes. Testing strategies, whether conversion analysis, as validated herein, or a combination of other approaches, must take into account the highly heterogeneous nature of mismatch repair mutations in colorectal cancer.

Acknowledgments

Funding/Support: GMP Genetics Inc sponsored this study and funded all the mutation analyses performed at GMP Genetics and City of Hope. This study was also funded by National Institutes of Health (NIH) grant U01 CA74800.

All other work was supported by funds from the National Cancer Institute and the NIH under RFA contract CA-95-011 and through cooperative agreements with the members of the Colon Cancer Family Registry and pricipal investigators.

Collaborating Centers: Australian Colorectal Cancer Family Registry (NIH grant UO1 CA097735), the USC Familial Colorectal Neoplasia Collaborative Group (NIH grant UO1 CA074799), Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (NIH grant UO1 CA074800), Ontario Registry for Studies of Familial Colorectal Cancer (NIH grant UO1 CA074783), Seattle Colorectal Cancer Family Registry (NIH grant UO1 CA074794), University of Hawaii Colorectal Cancer Family Registry (NIH grant UO1 CA074806), and University of California, Irvine Informatics Center (NIH grant UO1 CA078296).

REFERENCES

1. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96:261–268. [PubMed: 14970275]
2. Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on hereditary nonpolyposis colorectal cancer (ICG-HNPCC). Dis Colon Rectum 1991;34:424–425. [PubMed: 2022152]
3. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 1999;116:1453–1456. [PubMed: 10348829]
4. de la Chapelle A. William Allan Award address: inherited human diseases: victories, challenges, disappointments. Am J Hum Genet 2002;2003;72:236–240. [PubMed: 12596787]
5. Lynch HT, de le Chapelle A. Hereditary colorectal cancer. N Engl J Med 2003;348:919–932. [PubMed: 12621137]
6. International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer. Mutation database. [January 19, 2005]. Available at: http://www.insight-group.org.
7. Yan H, Papadopoulos N, Marra G, et al. Conversion of diploidy to haploidy. Nature 2000;403:723–724. [PubMed: 10693791]
8. Nakagawa H, Yan H, Lockman J, et al. Allele separation facilitates interpretation of potential splicing alterations and genomic rearrangements. Cancer Res 2002;62:4579–4582. [PubMed: 12183410]
9. Wagner A, Barrows A, Wijnen JT, et al. Molecular analysis of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterization of an American founder genomic deletion of the *MSH2* gene. Am J Hum Genet 2003;72:1088–1100. [PubMed: 12658575]
10. Lynch HT, Coronel SM, Okimoto R, et al. A founder mutation of the MSH2 gene and hereditary nonpolyposis colorectal cancer in the United States. JAMA 2004;291:718–724. [PubMed: 14871915]
11. Marra G, D'Atri S, Yan H, et al. Phenotypic analysis of hMSH2 mutations in mouse cells carrying human chromosomes. Cancer Res 2001;61:7719–7721. [PubMed: 11691782]
12. Liu T, Yan H, Kuismanen S, et al. The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. Cancer Res 2001;61:7798–7802. [PubMed: 11691795]
13. Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. J Clin Oncol 2002;20:1043–1048. [PubMed: 11844828]
14. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248–5257. [PubMed: 9823339]
15. Papadopoulos N, Leach FS, Kinzler KW, Vogel-stein B. Monoallelic mutation analysis (MAMA) for identifying germline mutations. Nat Genet 1995;11:99–102. [PubMed: 7550326]
16. Baudhuin LM, Mai M, French AJ, et al. Analysis of *hMLH1* and *hMSH2* gene dosage alterations in Lynch syndrome patients by novel methods. J Mol Diagn. In press.
17. Lynch HT, Coronel SM, Okimoto R, et al. A founder mutation of the *MSH2* gene and hereditary nonpolyposis colorectal cancer in the United States. JAMA 2004;291:718–724. [PubMed: 14871915]

18. Gille JJ, Hogervorst FB, Pals G, et al. Genomic deletions of *MSH2* and *MLH1* in colorectal cancer families detected by a novel mutation detection approach. Br J Cancer 2002;87:892–897. [PubMed: 12373605]

19. Nakagawa H, Hampel H, de la Chapelle A. Identification and characterization of genomic rearrangements of *MSH2* and *MLH1* in Lynch syndrome (HNPCC) by novel techniques. Hum Mutat 2003;22:258. [PubMed: 12938096]

20. Taylor CF, Charlton RS, Burn J, Sheridan E, Taylor GR. Genomic deletions in *MSH2* or *MLH1* are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA. Hum Mutat 2003;22:428–433. [PubMed: 14635101]

21. Baudhuin LM, Ming M, French AJ, et al. Analysis of *hMLH1* and *hMSH2* gene dosage alterations in Lynch syndrome patients by novel methods. J Molec Diagn. In press.

22. Liu B, Farrington SM, Petersen GM, et al. Genetic instability occurs in the majority of young patients with colorectal cancer. Nat Med 1995;1:348–352. [PubMed: 7585065]

23. Jass JR. Clinical significance of early-onset “sporadic” colorectal cancer with microsatellite instability. Dis Colon Rectum 2003;46:1305–1309. [PubMed: 14530666]

Table 1

Cancer History in Families With Defective Mismatch Repair

	Colorectal Cancer	Endometrial/Uterine Cancer	Other HNPCC-Related Cancers*
HNPCC cases (n = 64)			
No. per family, mean†	5.8	0.6	1.3
Age at onset, mean (range), y	48.0 (21-87)	45.4 (31-62)	53.1 (10-80)
HNPCC-like cases (n = 8)			
No. per family, mean†	1.3	0.8	0.8
Age at onset, mean (range), y	53.1 (10-80)	46.8 (43-51)	40.5 (42-66)
Case age <50 y (n = 17)			
No. per family, mean†	1.9	0.2	0.4
Age at onset, mean (range), y	49.9 (40-66)	41.3 (33-51)	53.9 (36-75)

Abbreviation: HNPCC, hereditary nonpolyposis colorectal cancer (Amsterdam 1 criteria).[2]

*Ovary, stomach, kidney, ureter, bladder, hepatobiliary, pancreas, small bowel, brain.

†Includes proband and first- and second-degree relatives.

Table 2

Mutations Identified by Conversion Analysis and DNA Sequencing in Cases With Defective Mismatch Repair*

	Mutation Type			
	MLH1	MSH2	MSH6	No./Total (%)
HNPCC cases (n = 64)				
Frameshift/nonsense/splice mutations†	21 (9)‡§	20 (14)§//	0	41/64 (64)
Genomic rearrangements/large genomic deletions	1 (0)	7 (0)	0	8/64 (13)
HNPCC-like cases (n = 8)				
Frameshift/nonsense/splice mutations	1 (1)	4 (2)§¶	0	5/8 (63)
Genomic rearrangements/large genomic deletions	0	2 (0)	0	2/8 (25)
Case age <50 y (n = 17)				
Frameshift/nonsense/splice mutations	0	3 (1)§¶	1 (1)	4/17 (24)
Genomic rearrangements/large genomic deletions	1 (0)	2 (0)	0	3/17 (18)

Abbreviation: HNPCC, hereditary nonpolyposis colorectal cancer (Amsterdam 1 criteria).[2]

*See “Results” section for pathogenic mutation criteria. Numbers in parentheses are mutations identified as deleterious by DNA sequencing.

†Two cases had the same *MLH1* splice-site mutation that correlated with loss of transcript expression.

‡Further pathogenic confirmation was required for 12 mutations.

§DNA sequencing identified the number of mutations in parentheses as likely deleterious.

//Further pathogenic confirmation was required for 5 mutations. The exon deletion mutation in ID No. 27 (see Table 3) was not identified by DNA sequencing.

¶Further pathogenic confirmation was required for 2 mutations.

Table 3

Colorectal Cancer Cases From HNPCC Amsterdam 1 Criteria Families With Defective Mismatch Repair: ID Nos. 1-42*

ID No.	Mutation	Mutation Consequence	Deleterious	Protein Expression
1	*MLH1* IVS08 + 3A>G	*MLH1* del exon 8 (exon skipping in cDNA)	Conversion analysis†	Loss of MLH1
2	*MLH1* 1852_1854delAAG	*MLH1* del K618 (in-frame deletion)	Insufficient data‡	Loss of MLH1
3	Mutation not identified by genomic DNA sequencing	*MSH2* del exons 1-8 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
4	*MLH1* 2199_2202insAACA	*MLH1* H733Xfs	Conversion analysis and DNA sequencing	Loss of MLH1
5	*MLH1* 1490_1491insC	*MLH1* R497Xfs	Conversion analysis and DNA sequencing	Loss of MLH1
6	*MLH1* VS07-2A>G	*MLH1* appears to result in unstable transcript (reported 3 times)§; loss of expression of affected *MLH1* allele	Conversion analysis†	Loss of MLH1
7	*MLH1* 1490_1901insC	*MLH1* R497Xfs	Conversion analysis and DNA sequencing	Loss of MLH1
8	*MSH2* IVS04 + 1G>A	*MSH2* del exon 4 (exon skipping in cDNA)	Conversion analysis†	Loss of MSH2 and MSH6
9	*MSH2* g.1566C>G *MSH2* c.1566C>G	*MSH2* Y522X; low expression of affected *MSH2* allele‖	Conversion analysis and DNA sequencing	Loss of MSH2 and MSH6
10	*MLH1* g.1459C>T and *MLH1* c.1459C>T	*MLH1* R487X	Conversion analysis and DNA sequencing	Loss of MLH1
11	*MLH1* IVS09 + 1G>A	*MLH1* del exons 9-10 (exon skipping in cDNA)	Conversion analysis†	Loss of MLH1
12	*MSH2* IVS15 + 5G>C	*MSH2* del exon 15 (exon skipping in cDNA); low expression of affected *MSH2* allele‖	Conversion analysis†	Loss of MSH2 and MSH6
20	MSH6 g.1273A>G MSH6 c.1273A>G	*MSH6* I425V (not reported)§	Insufficient data (reported as unclassified variant)¶	Loss of MLH1
21	*MSH2* g.1165C>T *MSH2* c.1165C>T	*MSH2* R389X	Conversion analysis and DNA sequencing	Loss of MSH2
22	Normal	No mutation detected	NA	Loss of MSH2
23	Normal	No mutation detected	NA	No IHC result#
24	*MSH2* g.1216C>T *MSH2* c.1216C>T	*MSH2* R406X; low expression of affected *MSH2* allele‖	Conversion analysis and DNA sequencing	Loss of MSH2
25	*MLH1* 1689_1690insA	*MLH1* L564Xfs (leading to skipping of exon 15)	Conversion analysis and DNA sequencing	No IHC result#
27	*MSH2* 136_164 del CACGGCGAGG ACGCGCTGCT GGCCGCCCG	*MSH2* H46_R55fs	Conversion analysis only	Loss of MSH2
28	*MLH1* g.793C>T *MLH1* c.793C>T	*MLH1* R265C (leading to skipping of exons 9-10 in cDNA); low expression of affected *MLH1* allele‖	Conversion analysis†	Loss of MLH1
29	*MLH1* g.793C>T *MLH1* c.793C>T	*MLH1* R265C (leading to skipping of exons 9-10 in cDNA); low expression of affected *MLH1* allele‖	Conversion analysis†	Loss of MLH1

ID No.	Mutation	Mutation Consequence	Deleterious	Protein Expression
30	*MSH2* g.2075G>T *MSH2* c.2075G>T	*MSH2* G692V (not previously reported)[§]	Insufficient data (reported as unclassified variant)[¶]	Loss of MSH2
31	*MLH1* 346delA	*MLH1* T116Xfs	Conversion analysis and DNA sequencing	Loss of MLH1
32	*MLH1* g.731G>A *MLH1* c.731G>A	*MLH1* G244D (reported once)[§]	Insufficient data (reported as unclassified variant)[¶]	Loss of MLH1
33	Mutation not identified by genomic DNA sequencing	*MSH2* del exon 8 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
34	*MLH1* IVS09 + 2T>C	*MLH1* del exons 9-10 (exon skipping in cDNA)	Conversion analysis[†]	Loss of MLH1
36	*MSH2* 2135_2136insT	*MSH2* V712Xfs leading to del exon 13 (exon skipping in cDNA); low expression of affected *MSH2* allele[//]	Conversion analysis and DNA sequencing	Loss of MSH2
37	*MLH1* IVS01 + 5G>C	*MLH1* c.116_117ins227nt (cryptic splice site in intron 1 resulting in longer cDNA)	Conversion analysis[†]	Loss of MLH1
41	*MSH2* g.2038C>T *MSH2* c.2038C>T	*MSH2* R680X; low expression of affected *MSH2* allele[//]	Conversion analysis and DNA sequencing	Loss of MSH2
42	*MLH1* g.199G>A *MLH1* c.199G>A	*MLH1* G67R (reported 6 times)[§]	Insufficient data (reported as unclassified variant)[¶]	Loss of MLH1 and PMS2

Abbreviations: HNPCC, hereditary nonpolyposis colorectal cancer; IHC, immunochemistry; NA, not applicable because no mutation was identified.

[*] Identification numbers are not consecutive. *MSH6* was not analyzed in most identification numbers. It was analyzed and classified as normal in identification numbers 22, 23, 30, 32, and 42; No. 20 had a MSH6 mutation at g.1273A>G and c.1273A>G.

[†] Mutation classified as pathogenic. Classification could not be made using genomic DNA sequencing alone.

[‡] Functional studies are required to confirm that in-frame deletion is deleterious. No other mutations were identified in *MLH1*, *MSH2*,or *MSH6*.

[§] Reported at http://www.insight-group.org.

[//] Refers to mRNA expression levels that were reduced (<90%) compared with control levels.

[¶] Additional studies are required to determine whether deleterious.

[#] Due to technical failure or assays not performed.

Table 4

Colorectal Cancer Cases From HNPCC Amsterdam 1 Criteria Families With Defective Mismatch Repair: ID Nos. 44-98*

ID No.	Mutation	Mutation Consequence	Deleterious	Protein Expression
44	*MLH1* g.350C>T *MLH1* c.350C>T	*MLH1* T117M (reported 10 times)†	Insufficient data (reported as unclassified variant)‡	Loss of MLH1
47	*MSH2* 154_155insG	*MSH2* L52Xfs	Conversion analysis and DNA sequencing	Loss of MSH2
48	*MSH2* 1593_1613 delAGTCCTTCGTAACAATAAAAA	*MSH2* K531_K537del (in-frame deletion)	Insufficient data§	MLH1 and MSH2 intact
50	*MLH1* g.883A>G *MLH1* c.883A>G	*MLH1* S295G (leading to skipping of exons 9-10 in cDNA)	Conversion analysis//	Loss of MLH1
51	*MLH1* g.350C>T *MLH1* c.350C>T	*MLH1* T117M (reported 10 times)†	Insufficient data (reported as unclassified variant)‡	Loss of MLH1
52	*MLH1* 1852_1854delAAG	*MLH1* del K618 (in-frame deletion)	Insufficient data§	Loss of MLH1
54	*MLH1* g.2074T>C *MLH1* c.2074T>C	*MLH1* S692P (not previously reported)†	Insufficient data (reported as unclassified variant)‡	MLH1 and MSH2 intact
55	*MSH2* 2237_2238insA	*MSH2* I746Xfs; low expression of affected *MSH2* allele¶	Conversion analysis and DNA sequencing	No IHC result#
56	Mutation not identified by genomic DNA sequencing	*MSH2* del exon 8 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
57	*MSH2* g.1660A>G *MSH2* c.1660A>G	*MSH2* del exon 10 (exon skipping in cDNA)	Conversion analysis//	Loss of MSH2
63	*MSH2* g.363T>G *MSH2* c.363T>G	*MSH2* Y121X; low expression of affected *MSH2* allele¶	Conversion analysis and DNA sequencing	Loss of MSH2
64	*MSH2* g.1373T>G *MSH2* c.1373T>G	*MSH2* L458X; low expression of affected *MSH2* allele¶	Conversion analysis and DNA sequencing	Loss of MSH2
66	*MSH2* 1705_1706delGA	*MSH2* E569Xfs	Conversion analysis and DNA sequencing	Loss of MSH2
67	*MLH1* g.1975C>T *MLH1* c.1975C>T	*MLH1* R659X leading to exon 17 skipping in cDNA	Conversion analysis and DNA sequencing	Loss of MLH1
68	*MSH2* IVS05 + 3A>T	*MSH2* del exon 5 (exon skipping in cDNA)	Conversion analysis//	Loss of MSH2
69	*MSH2* 1534_1543 delAAACTGGATT	*MSH2* K512Xfs; low expression of affected *MSH2* allele¶	Conversion analysis and DNA sequencing	Loss of MSH2
71	*MSH2* VS05 + 3A>T	*MSH2* del exon 5 (exon skipping in cDNA)	Conversion analysis//	Loss of MSH2
72	Mutation not identified by genomic DNA sequencing	*MLH1* del exons 2-6 genomic DNA; low expression of affected *MLH1* allele¶	Conversion analysis only	Loss of MLH1

ID No.	Mutation	Mutation Consequence	Deleterious	Protein Expression
75**	*MLH1* VS01 + 5G>C	*MLH1* c. 116_117ins227nt (cryptic splice site in intron 1 resulting in longer cDNA)	Conversion analysis//	Loss of *MLH1*
76	*MLH1* 503_504insA	*MLH1* N168Xfs	Conversion analysis and DNA sequencing	Loss of MLH1
77	*MSH2* g.1216C>T *MSH2* c.1216C>T	*MSH2* R406X	Conversion analysis and DNA sequencing	Loss of MSH2
78	*MLH1* g.113A>G *MLH1* c.113A>G	*MLH1* N38S (not previously reported)†	Insufficient data (reported as unclassified variant)‡	Loss of PMS2
79	*MLH1* 997_1000delAAGC	*MLH1* K33Xfs	Conversion analysis and DNA sequencing	Loss of MLH1
80	*MLH1* g.350C>G *MLH1* c.350C>G	*MLH1* T117R (reported twice)†; *MSH6* T764N (not previously reported)†	Insufficient data (reported as unclassified variant)//	No IHC result#
82	*MLH1* VS07-2A>G	*MLH1* appears to result in unstable transcript (reported 3 times)†; loss of expression of affected *MLH1* allele	Conversion analysis//	MLH1, MSH2, and MSH6 intact
84	*MLH1* VS07 + 5G>A	*MLH1* del exon 7 (exon skipping in cDNA)	Conversion analysis//	MLH1, MSH2, and MSH6 intact
85	Mutation not identified by genomic DNA sequencing	*MSH2* dup exon 7 in genomic DNA	Conversion analysis only	Loss of MSH2
92	*MSH2* 1222_1223insT	*MSH2* Y408Xfs	Conversion analysis and DNA sequencing	Loss of MSH2
93	*MLH1* IVS01-2A>G	*MLH1* c. 117_121delTTTAGfs at start of exon 2 in cDNA	Conversion analysis//	Loss of MLH1
94	Mutation not identified by genomic DNA sequencing	*MSH2* del exons 1-6 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
95	Normal	No mutation detected	NA	Loss of MSH2 and MSH6
97	Mutation not identified by genomic DNA sequencing	*MSH2* del exons 1-16 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
98	Mutation not identified by genomic DNA sequencing	*MSH2* del exons 4-16 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2

Abbreviations: HNPCC, hereditary nonpolyposis colorectal cancer; IHC, immunohistochemistry; NA, not applicable because no mutation was identified.

* Identification numbers are not consecutive. *MSH6* was not analyzed in most identification numbers. It was analyzed and classified as normal in identification numbers 44, 51, 54, and 95; No. 80 had an *MSH6* mutation at g.2291C>A and c.2291C>A.

† Reported at http://www.insight-group.org.

‡ Additional studies are required to determine whether deleterious.

§ Functional studies are required to confirm that in-frame deletion is deleterious. No other mutations were identified in *MLH1*, *MSH2*,or *MSH6*.

[//] Mutation classified as pathogenic. Classification could not be made using DNA sequencing alone.

[¶] Refers to mRNA expression levels that are reduced (<90%) compared with control levels.

[#] Due to technical failure or assays not performed.

[**] No data on microsatellite instability status.

Table 5

Eight Colorectal Cancer Cases From HNPCC-Like Familes With Defective Mismatch Repair*

ID No.	Mutation	Mutation Consequence	Deleterious	Protein Expression
38	Mutation not identified by genomic DNA sequencing	*MSH2* del exon 1 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
39	*MSH2* g.1738G>T *MSH2* c.1738G>T	*MSH2* E580X; low expression of affected *MSH2* allele†	Conversion analysis and DNA sequencing	Loss of MSH2
43	*MSH2* IVS05 + 3A>T	*MSH2* del exon 5 (exon skipping in cDNA)	Conversion analysis‡	Loss of MSH2 and MSH6
45	*MLH1* IVS11-8T>A	*MLH1* IVS11-8T>A had no effect on splicing; allele frequency 2.97% (reported once)§	NA	Loss of PMS2
65	Mutation not identified by genomic DNA sequencing	*MSH2* del exons 1-6 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
87	*MSH2* g.1034G>A *MSH2* c.1034G>A	*MSH2* W345X; loss of expression of affected *MSH2* allele	Conversion analysis and DNA sequencing	Loss of MSH2
89	*MSH2* IVS05 + 3A>T	*MSH2* del exon 5 (exon skipping in cDNA)	Conversion analysis‡	Loss of MSH2
91‖	*MLH1* 67delG	*MLH1* E23Xfs	Conversion analysis and DNA sequencing	Loss of MLH1

Abbreviations: HNPCC, hereditary nonpolyposis colorectal cancer; NA, not applicable because no mutation was identified.

**MSH6* was not analyzed in most identification numbers. It was analyzed and classified as normal in identification number 45.

†Refers to mRNA expression levels that are reduced (<90%) compared with control levels.

‡Mutation classified as pathogenic. Classification could not be made using genomic DNA sequencing alone.

§Reported at http://www.insight-group.org.

‖No data on microsatellite instability status.

Table 6

Seventeen Colorectal Cancer Cases With Defective Mismatch Repair Who Were Diagnosed Prior to Age 50 Years*

ID No.	Mutation	Mutation Consequence	Deleterious	Protein Expression
13	*MLH1* g.637G>A *MLH1* c.637G>A	*MLH1* V213M (reported 4 times)†	Insufficient data (reported as unclassified variant)‡	Loss of MLH1
14	Normal	No mutation detected	NA	Loss of MSH2 and MSH6
15	Normal	No mutation detected	NA	Loss of MLH1
16	Mutation not identified by genomic DNA sequencing	*MSH2* del exon 8 genomic DNA; low expression of affected *MSH2* allele§	Conversion analysis	Loss of MSH2
17//	*MSH2* 2680_2681insA	*MSH2* M896Xfs	Conversion analysis and DNA sequencing	Loss of MSH2 and MSH6
18	Normal	No mutation detected	NA	MLH1, MSH2, and MSH6 intact
19	Normal	No mutation detected	NA	Loss of MLH1
26	*MSH2* IVS9-2A>G	*MSH2* 1510_1511G; G504Xfs	Conversion analysis¶	Loss of MSH2
59	*MLH1* IVS11-8T>A	*MLH1*: no effect on splicing; allele frequency 2.97% (reported once)†	NA	No IHC result#
60	Normal	No mutation detected	NA	No IHC result#
62	Mutation not identified by genomic DNA sequencing	*MSH2* del exons 1-6 genomic DNA; loss of expression of affected *MSH2* allele	Conversion analysis only	Loss of MSH2
81//	*MSH2* 2576_2584delAATCGCAAG	*MSH2* Q859_E861del (in-frame deletion)	Insufficient data**	MLH1, MSH2, and MSH6 intact
83//	Mutation not identified by genomic DNA sequencing	*MLH1* del exons 16-19 genomic DNA *MLH1* IVS11-8T>A: no effect on splicing; allele frequency 2.97% (reported† once); loss of expression of affected *MLH1* allele	Conversion analysis only	MLH1, MSH2, and MSH6 intact
86//	*MLH1* g.2146G>A *MLH1* c.2146G>A	*MLH1* V716M (reported twice)†	Insufficient data (reported as unclassified variant)‡	No IHC result#
88//	*MSH6* 650_651insT	*MSH6* K218Xfs *MLH1* IVS11-8T>A: no effect on splicing; allele frequency 2.97% (reported once)†	Conversion analysis and DNA sequencing for 650_651insT	Loss of MSH6
90//	*MSH2* IVS05 + 3A>T	*MSH2* del exon 5 (exon skipping in cDNA)	Conversion analysis¶	No IHC result#
96	*MLH1* IVS11-8T>A	*MLH1* IVS11-8T>A: no effect on splicing; allele frequency 2.97% (reported once)†	NA	Loss of MSH2

Abbreviations: IHC, immunohistochemistry; NA, not applicable because no mutation was identified.

**MSH6* was not analyzed in identification numbers 16, 17, 26, 62, 81, 83, and 90. It was analyzed and classified as normal in identification numbers 13, 14, 15, 18, 19, 59, 60, 86, and 96; No. 88 had a MSH6 mutation in 650_651insT.

†Reported at http://www.insight-group.org.

‡Additional studies are required to determine whether deleterious.

§Refers to mRNA expression levels that are reduced (<90%) compared with control levels.

[//] No data on microsatellite instability status. Unable to reevaluate IHC staining in those cases with discordant mutation and IHC data.

[¶] Mutation classified as pathogenic. Classification could not be made using genomic DNA sequencing alone.

[#] Due to technical failure or assays not performed.

[**] Functional studies are required to confirm that in-frame deletion is deleterious. No other mutations were identified in *MLH1*, *MSH2*,or *MSH6*.